Gaussian State Amplification with Noisy State Observations

Bernd Bandemer
ITA Center, UC San Diego
La Jolla, CA 92093, USA
Email: bandemer@ucsd.edu

Chao Tian
AT&T Labs-Research
Florham Park, NJ 07932, USA
Email: tian@research.att.com

Shlomo Shamai (Shitz)
Dept. of EE, Technion–IIT
Technion City, Haifa 32000, Israel
Email: sshlomo@ee.technion.ac.il

***Abstract*—The problem of simultaneous message transmission and state amplification in a Gaussian channel with additive Gaussian state is studied when the sender has imperfect non-causal knowledge of the state sequence. Inner and outer bounds to the rate–state-distortion region are provided. The coding scheme underlying the inner bound combines analog signaling and Gelfand–Pinsker coding, where the latter deviates from the operating point of Costa's dirty paper coding.**

I. Introduction

Consider a Gaussian channel with additive Gaussian state in which the receiver simultaneously recovers a message communicated by the sender and estimates the state sequence. The sender facilitates this process by utilizing its own (possibly imperfect) knowledge of the state.

This problem was first investigated by Sutivong *et al.* [1] for the case when the sender has perfect knowledge of the state sequence before transmission begins. The authors show that the optimal rate–state-distortion tradeoff is achieved by dividing the available transmit power between analog state transmission and message transmission via dirty paper coding [2]. Subsequently, Kim *et al.* [3] considered the discrete memoryless version of this problem and characterized the tradeoff when the state reconstruction accuracy is measured by blockwise mutual information instead of quadratic distortion.

In this work, we are interested in the case in which the state observation at the sender is not perfect, but is encumbered by additive Gaussian state observation noise. A generalization of the problem in [1], this setting can be understood as modeling the original state amplification system with imperfect processing components. It also applies to the relay channel (see, for example, [4]), where the relay node attempts to amplify the primary transmitter's signal of which it has obtained a lossy description through another route.

Our setting contains several interesting extreme cases. When the state observation noise becomes negligible, our setting reverts to that of [1]. Conversely, the case when the state observation noise grows infinite, and thus the transmitter has no state knowledge, was studied in [5]. Pure state amplification without message transmission was considered in a previous work by the second author [6], where it was shown that an analog scheme with power control is optimal. By contrast, in the case of pure message transmission without state amplification, the optimal rate is achieved by dirty paper coding [2] with respect to the observable part of the state.

In this work, we consider the general tradeoff between message transmission rate and state amplification accuracy. We propose an inner bound and two outer bounds to the rate–state-distortion region. The inner bound is obtained by a hybrid scheme of analog state signaling and Gelfand–Pinsker encoding, where the Gelfand–Pinsker code in general does not coincide with dirty paper coding [2], but instead requires an optimized coefficient choice. The first outer bound is derived by generalizing the noise-partition approach in [6], while the second outer bound follows from careful analysis of the correlation structure in the problem.

In the following, we first provide a precise problem definition, before discussing the inner and outer bounds in Sections III and IV. We omit some details of the proofs for brevity's sake. Numerical examples and concluding remarks are given in Section V. Our mathematical notation follows [4].

II. Problem definition

Consider the state-dependent memoryless channel with input $X_i \in \mathbb{R}$ for $i = 1, 2, \ldots$ and output

$$Y_i = X_i + S_i + Z_i,$$

where $\{S_i\}$ and $\{Z_i\}$ are additive i.i.d. state and noise sequences, respectively, distributed according to $S_i \sim \mathrm{N}(0, Q)$ and $Z_i \sim \mathrm{N}(0, N)$. The channel input is subject to an average power constraint P. The sender has non-causal access to the noisy state observation

$$V_i = S_i + U_i,$$

where $\{U_i\}$ is an i.i.d. state observation noise sequence distributed according to $U_i \sim \mathrm{N}(0, \sigma_u^2)$. We assume that $\{S_i\}$, $\{Z_i\}$, and $\{U_i\}$ are independent.

The sender aims to communicate a message M at rate R reliably to the receiver. In addition to recovering the message, the receiver is also interested in obtaining an estimate $\hat{S}^n$ of the state sequence S^n, with state distortion measured by the

The work of S. Shamai (Shitz) was supported by the Israel Science Foundation (ISF), and the European Commission in the framework of the Network of Excellence in Wireless COMmunications NEWCOM#.

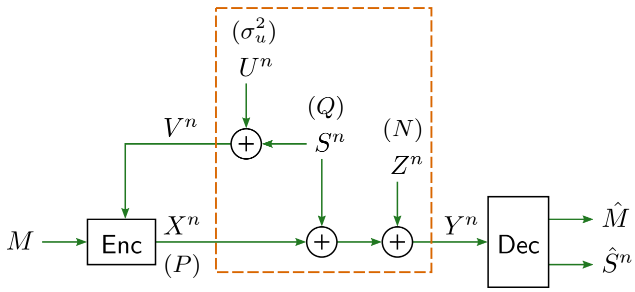

Figure 1. State amplification with noisy state observations.

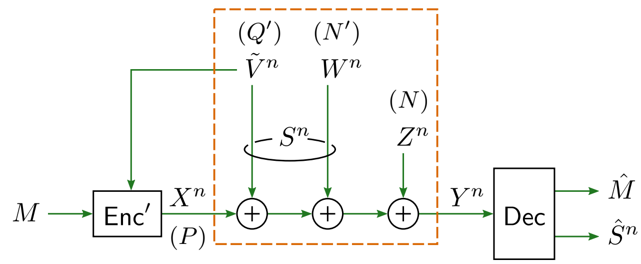

Figure 2. Equivalent system model.

average squared error and upper-bounded by D. The setup is depicted in Figure 1.

Formally, a $(2^{nR}, n)$ code for the Gaussian state amplification channel with noisy state observations consists of an encoder that maps a message $m \in \{1 : 2^{nR}\}$ and a state observation $v^n \in \mathbb{R}^n$ to a transmit sequence $x^n \in \mathbb{R}^n$ and a decoder that maps the received sequence $y^n \in \mathbb{R}^n$ to a message estimate $\hat{m} \in \{1 : 2^{nR}\}$ and a state estimate $\hat{s}^n \in \mathbb{R}^n$. We assume the message is random according to $M \sim \mathrm{Unif}\{1 : 2^{nR}\}$ and restrict our attention to codes that satisfy $(1/n)\sum_{i=1}^n \mathsf{E}[X_i^2] \le P$. A rate–state-distortion pair (R, D) is said to be achievable if there exists a sequence of $(2^{nR}, n)$ codes such that $\lim_{n\to\infty} \mathsf{P}\{\hat{M} \ne M\} = 0$ and $\limsup_{n\to\infty} (1/n)\sum_{i=1}^n \mathsf{E}(S_i - \hat{S}_i)^2 \le D$. We are interested in characterizing the rate–distortion region, i.e., the closure of the set of achievable (R, D) pairs.

Before we begin our discussion, it is helpful to reformulate the problem as follows. Since the triple (S_i, U_i, V_i) is jointly Gaussian, we can equivalently write

$$S_i = \tilde{V}_i + W_i,$$

with $\tilde{V}_i = Q/(Q + \sigma_u^2)V_i \sim \mathrm{N}(0, Q')$ and $W_i \sim \mathrm{N}(0, N')$. Here, $\{\tilde{V}_i\}$ and $\{W_i\}$ are independent i.i.d. Gaussian sequences with variances

$$Q' = \frac{Q^2}{Q + \sigma_u^2},$$
$$N' = \frac{Q\sigma_u^2}{Q + \sigma_u^2}.$$

This equivalent system model is depicted in Figure 2. Note that the channel state S^n is decomposed into a perfectly observable part $\tilde{V}^n$ and a completely unobservable part W^n. However, the receiver still aims to estimate the entire state $\tilde{V}^n + W^n$.

III. Inner Bound

For a constant $\beta \in [0, 1]$, let

$$g = \sqrt{(1-\beta)P/Q'},$$
$$r = \begin{bmatrix} (1+g)Q' + N' \\ \alpha(1+g)Q' \end{bmatrix}, \tag{1}$$
$$\Sigma = \begin{bmatrix} (1+g)^2Q' + \beta P + N' + N & \beta P + \alpha(1+g)^2Q' \\ \beta P + \alpha(1+g)^2Q' & \beta P + \alpha^2(1+g)^2Q' \end{bmatrix}. \tag{2}$$

Then we have the following inner bound to the rate–state-distortion region.

Theorem 1. The rate–state-distortion (R, D) is achievable if

$$R < \tfrac{1}{2}\log \frac{\beta P(\beta P + (1+g)^2Q' + N' + N)}{(N'+N)(\beta P + \alpha^2(1+g)^2Q') + (1-\alpha)^2\beta(1+g)^2PQ'}, \tag{3}$$
$$D \ge Q' + N' - r^{\mathrm{T}}\Sigma^{-1}r, \tag{4}$$

for some $\alpha \in \mathbb{R}_+$, $\beta \in [0, 1]$.

Proof: Consider the following coding scheme, in which the transmit signal consists of a scaled version of the channel state sequence and a Gelfand–Pinsker codeword matched to the amplified state sequence. More formally, we generate a codebook as follows.

Codebook generation: Fix α and β. Let $\tilde{X} \sim \mathrm{N}(0, \beta P)$ be independent of the state $\tilde{V}$, and $U = \tilde{X} + \alpha(1+g)\tilde{V}$. Define an auxiliary rate $\tilde{R} \ge R$. For each message $m \in \{1 : 2^{nR}\}$, generate a subcodebook $\mathcal{C}(m)$ consisting of $2^{n(\tilde{R}-R)}$ sequences $u^n(m, l)$, for $l \in \{1 : 2^{n(\tilde{R}-R)}\}$, each independently generated according to $\prod_{i=1}^n p_U(u_i)$.

Encoding: Fix $\varepsilon' > 0$. To communicate message m given the state sequence $\tilde{v}^n$, the sender finds an index l such that $(u^n(m, l), \tilde{v}^n) \in \mathcal{T}_{\varepsilon'}^{(n)}(U, \tilde{V})$, and transmits

$$x_i = g\tilde{v}_i + \tilde{x}_i, \quad \text{for } i \in \{1 : n\},$$

where $\tilde{x}_i = u_i(m, l) - \alpha(1+g)\tilde{v}_i$. Observe that by construction, the sequence $\tilde{x}^n$ is jointly typical with the state sequence $\tilde{v}^n$, that is, symbolwise pairs $(\tilde{x}_i, \tilde{v}_i)$ are asymptotically independent. Therefore, the average power constraint P is satisfied.

Decoding and analysis of the probability of error: Observe y^n. Let $\varepsilon > \varepsilon'$. Declare that message $\hat{m}$ has been sent if $(\hat{m}, \hat{l})$ is the unique index pair such that

$$(u^n(\hat{m}, \hat{l}), y^n) \in \mathcal{T}_{\varepsilon}^{(n)}.$$

By the result of Gelfand–Pinsker for channels with state [7], the probability of decoding error vanishes as $n \to \infty$ if

$$R < I(U; Y) - I(U; \tilde{V}).$$

To evaluate the terms, recall that

$$Y = (1+g)\tilde{V} + \tilde{X} + W + Z, \tag{5}$$
$$U = \tilde{X} + \alpha(1+g)\tilde{V}, \tag{6}$$

where $\tilde{V}$, $\tilde{X}$, W, and Z are independent Gaussians of variances Q', βP, N', and N. It is not hard to see that the rate condition evaluates to (3).

Estimation and analysis of estimation error: For each symbol time $i \in \{1:n\}$, construct the best MSE estimate of s_i given the observations y_i and $u_i(\hat{m}, \hat{l})$, where $\hat{m}$ and $\hat{l}$ are the decoded message and subcode index, respectively. Omitting the time index, recall (5), (6) and

$$S = \tilde{V} + W, \tag{7}$$

Hence the joint distribution of (S, Y, U) is

$$\begin{bmatrix} S \\ Y \\ U \end{bmatrix} \sim \mathrm{N}\left(0, \begin{bmatrix} Q' + N' & r^{\mathrm{T}} \\ r & \Sigma \end{bmatrix}\right),$$

where r and Σ are shorthand for the cross-correlation vector between S and $[Y, U]$, and the autocorrelation matrix of $[Y, U]$, respectively, and are given in (1) and (2). The best estimate is

$$\hat{s}_i = r^{\mathrm{T}} \Sigma^{-1} \begin{bmatrix} y_i \\ u_i(\hat{m}, \hat{l}) \end{bmatrix}, \tag{8}$$

and the mean square error satisfies (4). This concludes the proof of Theorem 1. ■

Remark 1. Recall that in Costa's dirty paper coding [2], the parameter α is chosen as

$$\alpha = \frac{\beta P}{\beta P + N}. \tag{9}$$

It turns out that varying α as in the theorem achieves a larger inner bound in our setting.

Remark 2. The estimate $\hat{s}_i$ in (8) is independent of u_i when the second component of $r^{\mathrm{T}} \Sigma^{-1}$ is zero,

$$\begin{bmatrix} (1+g)Q' + N' \\ \alpha(1+g)Q' \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} -\beta P - \alpha(1+g)^2 Q') \\ (1+g)^2 Q' + \beta P + N' + N \end{bmatrix} = 0,$$

which occurs when

$$\alpha = \frac{\beta P((1+g)Q' + N')}{(1+g)Q'(\beta P + N - gN')}. \tag{10}$$

At all other values of α, u_i is useful in estimating s_i. As a side note, observe that if there is no state observation noise (when $N' = 0$), (10) reduces to (9).

Remark 3. Numerical computations indicate that the inner bound in Theorem 1 cannot be improved by (1) reducing the gain g, (2) using part of the message rate to send a digital description of $\tilde{V}^n$, or (3) diverting a fraction of power to send an additional message by superposition coding.

IV. Outer bounds

We present two outer bounds to the rate–state-distortion region. Recall $\lambda = Q/(Q + \sigma_u^2)$.

Theorem 2. If a rate–state-distortion pair (R, D) is achievable, then for all $\bar{N} \in [0, N]$, there exists an $\bar{r} \in [0, \sqrt{P(Q + \sigma_u^2)}]$ such that

$$R \le \frac{1}{2} \log \frac{(Q + \sigma_u^2)(P + N + E_S) - \bar{r}^2}{(Q + \sigma_u^2)(N - \bar{N})}, \tag{11}$$

$$D \ge \left(1 + 2^{2R} \cdot \frac{Q(\bar{N} + \sigma_u^2)(N - \bar{N})}{\bar{N}\sigma_u^2(P + Q + N + 2\lambda\bar{r})}\right) E_S, \tag{12}$$

where

$$E_S = \frac{Q\bar{N}\sigma_u^2}{Q\bar{N} + Q\sigma_u^2 + \bar{N}\sigma_u^2}.$$

Proof: Following [6], let us divide the noise Z_i into two independent components $\bar{Z}_i$ and $\bar{\bar{Z}}_i$, with

$$\bar{Z}_i \sim \mathrm{N}(0, \bar{N}),$$
$$\bar{\bar{Z}}_i \sim \mathrm{N}(0, N - \bar{N}).$$

Note that E_S, as defined in the theorem, denotes the mean square error of the best linear estimator of S_i given $V_i = S_i + U_i$ and $S_i + \bar{Z}_i$. Define the quantity

$$\Delta = \tfrac{1}{n} I(V^n, S^n + \bar{Z}^n; Y^n).$$

Using results from the remote source coding problem in rate–distortion theory, it was shown in [6] that the information measure Δ and the quadratic distortion D are related as

$$\Delta \ge \tfrac{1}{2} \log\left(\frac{Q(\bar{N} + \sigma_u^2)}{\bar{N}\sigma_u^2(D/E_S - 1)}\right),$$

or equivalently,

$$D \ge \left(1 + 2^{-2\Delta} \cdot \frac{Q(\bar{N} + \sigma_u^2)}{\bar{N}\sigma_u^2}\right) E_S. \tag{13}$$

This allows us to translate upper bounds on Δ to lower bounds on D. Next, we obtain an outer bound on the achievable (R, Δ) region. First, it follows from Fano's inequality that

$$\begin{aligned} nR &= H(M) \\ &= H(M \mid V^n, S^n + \bar{Z}^n) \\ &\le I(M; Y^n \mid V^n, S^n + \bar{Z}^n) + n\varepsilon_n \\ &= h(Y^n \mid V^n, S^n + \bar{Z}^n) - h(\bar{\bar{Z}}^n) + n\varepsilon_n. \end{aligned} \tag{14}$$

Let

$$\bar{r} = \frac{1}{n} \sum_{i=1}^{n} |\mathsf{E}(X_i V_i)|. \tag{15}$$

Using covariance matrices to bound differential entropies, it can be shown that

$$h(Y^n) \le \tfrac{n}{2} \log\bigl(2\pi e(P + Q + N + 2\lambda\bar{r})\bigr), \tag{16}$$

$$h(Y^n \mid V^n, S^n + \bar{Z}^n) \le \tfrac{n}{2} \log\bigl(2\pi e\bigl(P + N + E_S - \bar{r}^2/(Q + \sigma_u^2)\bigr)\bigr). \tag{17}$$

Substituting these inequalities and the definition of Δ into (14), we conclude that (R, Δ) must satisfy (11) and

$$\Delta \le \tfrac{1}{2} \log \frac{P + Q + N + 2\lambda\bar{r}}{N - \bar{N}} - R$$

as $n \to \infty$. Using the last inequality with (13), we obtain (12).

Finally, note that $\bar{r} \in [0, \sqrt{P(Q+\sigma_u^2)}]$ since

$$\begin{aligned} \bar{r} &\overset{\text{(a)}}{\leq} \sqrt{\frac{1}{n}\sum_{i=1}^{n} \mathsf{E}(X_i V_i)^2} \\ &\overset{\text{(b)}}{\leq} \sqrt{\frac{1}{n}\sum_{i=1}^{n} \mathsf{E}(X_i^2)(Q+\sigma_u^2)} \\ &\overset{\text{(c)}}{\leq} \sqrt{P(Q+\sigma_u^2)}, \end{aligned}$$

where (a) follows from the generalized mean inequality, (b) follows from the Cauchy–Schwarz inequality, and (c) follows from the power constraint. This concludes the proof of the outer bound in Theorem 2. ∎

In order to state the second outer bound, we define the function f as

$$f(x) = (\sqrt{x} - \sqrt{N'/Q'}\sqrt{Q'-x})_+^2,$$

where x_+ denotes the positive part $\max\{x, 0\}$. It can be shown that f is convex and non-decreasing.

Theorem 3. If a rate–distortion pair (R, D) is achievable, then it must satisfy

$$R \leq \frac{1}{2}\log \frac{\sigma_u^2(N+P+Q) + Q(N+P) - \bar{r}^2}{(Q+\sigma_u^2)(N+N')}, \tag{18}$$

$$D \geq f\left(\frac{Q'(N+N')}{P+Q+N+2\lambda\bar{r}}\, 2^{2R}\right) \tag{19}$$

for some $\bar{r} \in [0, \sqrt{P(Q+\sigma_u^2)}]$.

Proof: Let $\Delta' = \frac{1}{n}I(\tilde{V}^n; Y^n)$. To relate Δ' and the distortion D, note that

$$\begin{aligned} \Delta' &\overset{\text{(a)}}{\geq} \tfrac{1}{n} I(\tilde{V}^n; \hat{S}^n) \\ &= \tfrac{1}{n} h(\tilde{V}^n) - \tfrac{1}{n} h(\tilde{V}^n \,|\, \hat{S}^n) \\ &\geq \tfrac{1}{n} h(\tilde{V}^n) - \tfrac{1}{n}\sum_{i=1}^{n} h(\tilde{V}_i \,|\, \hat{S}_i) \\ &\geq \tfrac{1}{n} h(\tilde{V}^n) - \tfrac{1}{2n}\sum_{i=1}^{n} \log(2\pi e \operatorname{Var}(\tilde{V}_i \,|\, \hat{S}_i)) \\ &\overset{\text{(b)}}{\geq} \tfrac{1}{n} h(\tilde{V}^n) - \tfrac{1}{2}\log\Big(2\pi e \cdot \tfrac{1}{n}\sum_{i=1}^{n} \operatorname{Var}(\tilde{V}_i \,|\, \hat{S}_i)\Big) \\ &\overset{\text{(c)}}{\geq} \tfrac{1}{n} h(\tilde{V}^n) - \tfrac{1}{2}\log\Big(2\pi e \cdot \tfrac{1}{n}\sum_{i=1}^{n} D_i'\Big) \\ &= \tfrac{1}{2}\log(Q') - \tfrac{1}{2}\log\Big(\tfrac{1}{n}\sum_{i=1}^{n} D_i'\Big), \end{aligned}$$

where (a) follows from the data processing inequality, (b) follows from Jensen's inequality and (c) follows from defining D_i' to be the mean square error of best *linear* estimator of $\tilde{V}_i$ given $\hat{S}_i$. Hence

$$\tfrac{1}{n}\sum_{i=1}^{n} D_i' \geq Q' \cdot 2^{-2\Delta'}. \tag{20}$$

Let $D_i = \operatorname{Var}(S_i - \hat{S}_i)$ be the mean square instantaneous state estimation error at the receiver at time i. In Appendix A, we use the Cauchy–Schwarz inequality to show

$$D_i \geq f(D_i').$$

Thus,

$$\begin{aligned} D = \tfrac{1}{n}\sum_{i=1}^{n} D_i &\geq \tfrac{1}{n}\sum_{i=1}^{n} f(D_i') \\ &\overset{\text{(a)}}{\geq} f\left(\tfrac{1}{n}\sum_{i=1}^{n} D_i'\right) \\ &\overset{\text{(b)}}{\geq} f(Q' \cdot 2^{-2\Delta'}), \end{aligned} \tag{21}$$

where (a) follows from Jensen's inequality and (b) follows from (20).

Next, we obtain an outer bound on the achievable (R, Δ') region. From Fano's inequality, we have

$$\begin{aligned} nR &= H(M \,|\, \tilde{V}^n) \\ &\leq I(M; Y^n \,|\, \tilde{V}^n) + n\varepsilon_n \\ &= h(Y^n \,|\, \tilde{V}^n) - \underbrace{h(Y^n \,|\, \tilde{V}^n, M)}_{= \frac{n}{2}\log(2\pi e(N'+N))} + n\varepsilon_n. \end{aligned} \tag{22}$$

Reusing definition (15), and using covariance matrices to bound differential entropies, it can be shown that

$$h(Y^n \,|\, \tilde{V}^n) \leq \tfrac{n}{2}\log\Big(2\pi e \frac{\sigma_u^2(N+P+Q) + Q(N+P) - \bar{r}^2}{Q+\sigma_u^2}\Big).$$

Substituting this inequality, bound (16) and the definition of Δ' into (22), we obtain (18) and

$$\Delta' \leq \tfrac{1}{2}\log\frac{P+Q+N+2\lambda\bar{r}}{N+N'} - R$$

as $n \to \infty$. Using the last inequality with (21) yields (19), which concludes the proof of Theorem 3. ∎

V. Numerical examples and Concluding remarks

Two examples for the inner bound of Theorem 1 and the outer bounds of Theorems 2 and 3 are depicted in Figures 3 and 4. In the example in Figure 4, the transmit power P is sufficiently large to permit a nonzero rate at the minimum distortion value. This is the case where the optimal transmission scheme without a message ($R = 0$) as discussed in [6] uses less than the full transmit power.

We observe that the outer bounds complement each other, but exhibit a nonnegligible gap from the inner bound. To close this gap, new outer bounding techniques will be necessary, the study of which is the subject of future research.

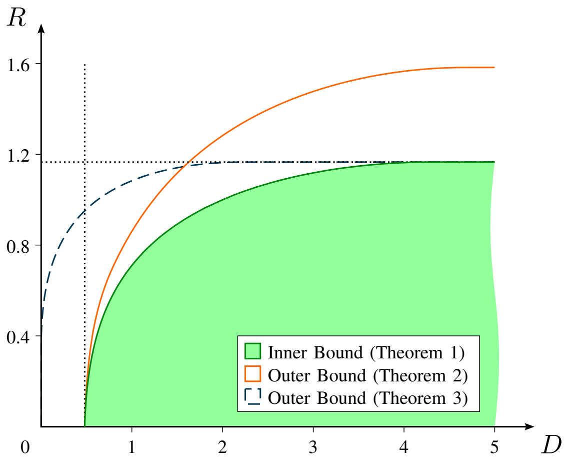

Figure 3. Bounds to the rate–estimation-error region, parameters $Q = 10$, $N = 1$, $\sigma_u^2 = 1$, and $P = 7.7$ (i.e., $Q' = 9.09$, $N' = 0.91$).

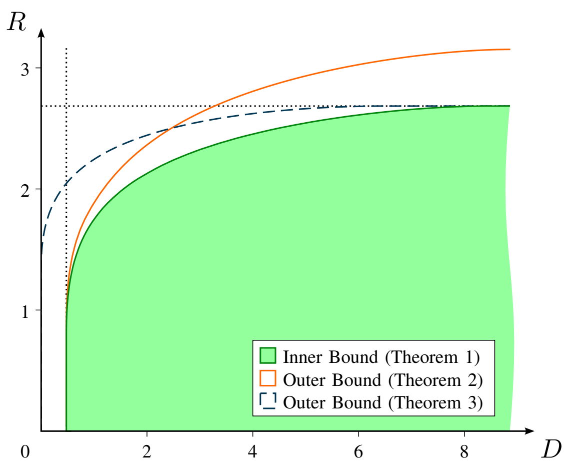

Figure 4. Bounds to the rate–estimation-error regions, parameters $Q = 10$, $N = 1$, $\sigma_u^2 = 1$, and $P = 77$ (i.e., $Q' = 9.09$, $N' = 0.91$).

Appendix

A. Relation between D_i and D_i'

Let $E_i = S_i - \hat{S}_i$ and recall $D_i = \mathrm{Var}(S_i - \hat{S}_i) = \mathrm{Var}(E_i)$. Likewise, let $\hat{V}_i$ be the best linear estimator of $\tilde{V}_i$ given $\hat{S}_i$, define the corresponding estimation error as $F_i = \tilde{V}_i - \hat{V}_i$ and recall $D_i' = \mathrm{Var}(F_i)$.

By the orthogonality principle, we have $\mathsf{E}(F_i \hat{V}_i) = 0$, and using $\mathsf{E}(\tilde{V}_i^2) = Q'$, it follows that

$$\mathsf{E}(\tilde{V}_i \hat{V}_i) = \mathsf{E}(\hat{V}_i^2) = Q' - D_i'.$$

Since $\hat{V}_i$ is a scaled version of $\hat{S}_i$, the variance D_i is lower bounded by the mean square error of the best linear estimator of S_i given $\hat{V}_i$, namely

$$D_i \geq Q' + N' - \frac{\mathsf{E}(\hat{V}_i S_i)^2}{Q' - D_i'}. \tag{23}$$

To evaluate the expectation, consider

$$\hat{V}_i = \frac{Q' - D_i'}{Q'} \tilde{V}_i + F_i',$$

where the first term in the sum is the best linear estimator of $\hat{V}_i$ given $\tilde{V}_i$, and F_i' is the corresponding estimation error. Thus, $\tilde{V}_i$ and F_i' are uncorrelated and

$$\mathrm{Var}(F_i') = Q' - D_i' - \frac{(Q' - D_i')^2}{Q'} = \frac{D_i'(Q' - D_i')}{Q'}.$$

Furthermore, recall

$$S_i = \tilde{V}_i + W_i,$$

where $\tilde{V}_i$ and W_i are independent and $\mathrm{Var}(W_i) = N'$. Thus,

$$\begin{aligned} \mathsf{E}(\hat{V}_i S_i) &= \mathsf{E}\left(\left(\frac{Q' - D_i'}{Q'} \tilde{V}_i + F_i' \right) (\tilde{V}_i + W_i) \right) \\ &= Q' - D_i' + \mathsf{E}(F_i' W_i). \end{aligned} \tag{24}$$

By the Cauchy–Schwarz inequality,

$$\begin{aligned} |\mathsf{E}(F_i' W_i)| &\leq \sqrt{\mathsf{E}(F_i'^2)\,\mathsf{E}(W_i^2)} \\ &= \sqrt{\frac{N' D_i'(Q' - D_i')}{Q'}}. \end{aligned}$$

Substituting back into (24) and using $D_i' \leq Q'$, this implies

$$|\mathsf{E}(\hat{V}_i S_i)| \leq Q' - D_i' + \sqrt{\frac{N' D_i'(Q' - D_i')}{Q'}}.$$

Further, substituting back into (23) yields

$$\begin{aligned} D_i &\geq Q' + N' - \frac{\left(Q' - D_i' + \sqrt{\frac{N' D_i'(Q' - D_i')}{Q'}} \right)^2}{Q' - D_i'} \\ &= \left(\sqrt{D_i'} - \sqrt{N'/Q'}\sqrt{Q' - D_i'} \right)^2 \\ &\geq \left(\sqrt{D_i'} - \sqrt{N'/Q'}\sqrt{Q' - D_i'} \right)_+^2 \\ &= f(D_i'), \end{aligned}$$

which concludes the proof.

References

[1] A. Sutivong, M. Chiang, T. M. Cover, and Y.-H. Kim, "Channel capacity and state estimation for state-dependent Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1486–1495, Apr. 2005.

[2] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, 1983.

[3] Y.-H. Kim, A. Sutivong, and T. M. Cover, "State amplification," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1850–1859, May 2008.

[4] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[5] W. Zhang, S. Vedantam, and U. Mitra, "Joint transmission and state estimation: A constrained channel coding approach," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 7084–7095, Oct. 2011.

[6] C. Tian, "Amplification of the hidden Gaussian channel states," in *Proceedings of ISIT*, Boston, MA, Jul. 2012.

[7] S. I. Gelfand and M. S. Pinsker, "Coding for channel with random parameters," *Probl. Control Inf. Theory*, vol. 9, no. 1, pp. 19–31, 1980.